Comments from JABA Readers

REAPPRAISAL OF MOMENTARY TIME SAMPLING AND PARTIAL-INTERVAL RECORDING

As a result of computer simulations of momentary time sampling (MTS) and partial-interval recording (PIR) for estimating known parameters, Harrop and Daniels (1986) recommended the use of PIR rather than MTS, even though their data indicated that MTS provides accurate estimates of absolute behavior duration and does not suffer from systematic errors, and that PIR estimates of duration are systematically positively biased. Their preference for PIR was based on their conclusion that PIR is more sensitive to relative changes although it underestimates change with high-rate behaviors. Sampling theory explains the results of the simulations but leads us to the conclusion that for the estimation of duration as defined by Harrop and Daniels, MTS, not PIR, should be the sampling mode of choice.

The explanation of why "MTS (but not PIR) provides accurate average estimates of absolute duration" (Harrop & Daniels, 1986, p. 76) is inherent in the definitions of the two sampling modes. The ability of both systems to reflect the true total duration of a behavior is, in part, a function of the proportion of sampling intervals during which the behavior either continued during the entire sample interval or never occurred during the entire interval. The basic source of error in both systems is the proportion of intervals in which behavior occurs during part of an interval but not all of it (mixed intervals). Longer emitted bouts of behavior will produce more intervals that contain all behavior and therefore will be correctly recorded by either MTS or PIR, whereas shorter bouts of behavior will produce more mixed intervals that are subject to error. The same is true of long versus short bouts of interresponse time (IRT).

By definition, PIR will record all mixed intervals as an occurrence unless the behavior occurs during a pause for recording. Therefore PIR will always overestimate duration. In the Harrop and Daniels (1986) data, overestimates as high as 900% are seen.

The expected outcomes of MTS can be predicted by sampling theory, because observing every nth second is the same as drawing a random sample of $1/n$th the size of the population.

Harrop and Daniels' (1986) second conclusion was that "MTS . . . appears not to suffer from systematic error in estimating relative change" (p. 76). It has been shown mathematically (Suen & Ary, 1989) that by definition MTS errors are unbiased; therefore, MTS (but not PIR) will produce unbiased estimates of relative change.

"PIR underestimates the magnitude of change with high rate behaviors" (p. 76) was Harrop and Daniels' (1986) third conclusion. The difference between MTS and PIR outcomes is due to the fact that MTS counts only a fraction of the mixed intervals as occurrences, whereas PIR counts all of them as behavior; this produces systematic overestimates of the percentage of total duration. Therefore, the difference between the PIR estimate and 100% is always less than the difference between percentage of true duration and 100%. The ability of PIR estimates to reflect true change becomes increasingly restricted as the PIR estimates approach 100%. This phenomenon is most pronounced when frequent behavior bouts of short duration produce a high proportion of mixed intervals and is least pronounced when long bouts of behavior and/or long interresponse times reduce the proportion of mixed intervals. Because the errors in MTS are unbiased, MTS duration estimates have a linear relation with true duration and provide unbiased estimates of changes in behavior under all conditions.

Harrop and Daniels (1986) also concluded that "estimates of absolute rate are inaccurate with both

methods" (p. 76). If we define rate as the number of times a behavior is initiated or how many bouts of behavior there are in a session, the proportion of MTS intervals scoring behavior has *no* inherent relationship with rate. If a behavior occurs for 360 s in a 1-hr session, MTS will, on average, record that it occurred 10% of the time whether the behavior was in one bout of 360 s, six 1-min bouts, 36 10-s bouts, or 360 1-s bouts. The only difference is that as the number of bouts increases, the number of mixed intervals increases and, therefore, the error increases.

For any given true total duration, MTS can be expected to record behavior in the same number of intervals whether that behavior is massed in one bout or is distributed among many bouts. However, if true duration remains constant, as the number of bouts increases, the number of mixed intervals increases and the number of intervals recorded as behavior occurrence increases with PIR but not with MTS. This is why Rhine and Linville (1980) found correlations between MTS results and frequency of emission.

PIR outcomes are inherently influenced by the number of bouts. Although there is a mathematical relationship between PIR outcomes and number of bouts, the relationship is curvilinear and practically useless in estimating how often the behavior was initiated.

Their final conclusion, that "PIR is more sensitive than MTS in detecting relative change in behavioral level (rate or total duration)" (p. 76), was apparently based on an analysis that recorded how many seconds were counted by PIR in each interval. This information is not available in practice. When Harrop and Daniels (1986) compared the outcomes of PIR and MTS when both score either 0 or 1, as is the case in actual practice, they found that MTS did not suffer from systematic error in estimating change but PIR underestimated the magnitude of change with high-rate behaviors. Therefore, in actual practice MTS is more sensitive in detecting relative changes in behavioral level than is PIR.

A mathematical explanation of the Harrop and Daniels (1986) study explains why they reached the conclusions they did and lead us to concur with most of them. However, their last sentence, "Because, however, the systematic error produced by PIR is always in the direction of providing a conservative estimate of change, researchers and practitioners may consider that this error is a price worth paying for the greater sensitivity of the method" (pp. 76–77), is apparently based on an analysis of data never available in practice. Our analysis of the same data leads us to concur with Powell (1984) in concluding that when duration (or other parameters involving duration in a ratio, proportion, or transformation) is the dimension of interest, MTS has much to recommend it and PIR has nothing to recommend it.

REFERENCES

Harrop, A., & Daniels, M. (1986). Methods of time sampling: A reappraisal of momentary time sampling and partial-interval recording. *Journal of Applied Behavior Analysis,* **19,** 73–77.

Powell, J. (1984). On the misrepresentation of behavioral realities by a widely practiced direct observation procedure: Partial interval (one-zero) sampling. *Behavioral Assessment,* **6,** 209–219.

Rhine, R., & Linville, A. (1980). Properties of one-zero scores in observational studies of primate social behavior: The effect of assumptions on empirical analysis. *Primates,* **21,** 11–112.

Suen, H. K., & Ary, D. (1989). *Analyzing quantitative behavioral observation data.* Hillsdale, NJ: Erlbaum.

Hoi K. Suen
PENNSYLVANIA STATE UNIVERSITY
Donald Ary and Wesley Covalt
NORTHERN ILLINOIS UNIVERSITY

Received July 30, 1990
Initial editorial decision March 13, 1991
Revisions received May 22, 1991
Final acceptance June 26, 1991
Action Editor, Terry J. Page